

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR LETTERS PATENT

**AN ARCHITECTURE AND RELATED METHODS FOR
STREAMING MEDIA CONTENT THROUGH
HETEROGENEOUS NETWORKS**

Inventor(s):

Qian Zhang

Wenwu Zhu

Ya-Qin Zhang

Fan Yang

ATTORNEY'S DOCKET NO. MS1-716US

05848705-050204
T.02050-90284860

1 **TECHNICAL FIELD**

2 This invention generally relates to an architecture and techniques for
3 transmitting content through heterogeneous networks.
4

5 **BACKGROUND**

6 With recent improvements in processor, storage and networking
7 technologies, today's wireless communications devices take many different forms,
8 and offer a number of innovative features. Examples of this new class of wireless
9 communication devices include personal digital assistant's (or, PDA's) fitted with a
10 wireless communications interface, two-way paging devices, digital
11 communications devices, and third-generation (or, 3G) personal communicators.
12 Gone are the days when such devices are merely used to support verbal
13 communication. Today, consumers are demanding that such devices be
14 multifunctional, enabling a user to receive and/or retrieve email, to send/receive
15 text messages, and to access text-based content from the Internet (e.g., stock
16 quotes, flight arrival/departure information, etc.). Despite the recent innovation in
17 the wireless communication space, consumers are demanding even greater
18 application resources and functionality for their portable communications device.
19 For example, despite recent advances in the functionality and capability of wireless
20 access devices, no good solution exists for streaming media content over
21 heterogeneous networks (e.g., a media streaming path comprising a wired network
22 component and wireless network component).

23 Those skilled in the art will appreciate that streaming media content over a
24 heterogeneous network comprised of a wired component and a wireless component
25 is a very challenging task. One reason is that the transport layer protocol
associated with each of the disparate components was developed independent of

one another and, therefore, without regard for many of the design challenges of the other. For example, the control algorithm of a typical wireline transport layer protocol (e.g., TCP, UDP, etc.) assumes that any degradation in transmission quality is due to congestion problems in one or more of the network elements (e.g., router, switch, hub, etc.). Accordingly, conventional transport layer protocols will iteratively reduce the transmission rate until transmission quality is improved. However, in heterogeneous networks that include a wireless network component, the degradation in transmission quality may have nothing to do with congestion on the wireline component of the communication channel. That is, the degradation in transmission quality, measured in terms of a bit-error rate (or, BER), may well be caused by a fading condition, a shadowing effect, multi-path sources, etc. in the wireless communication channel. It will be appreciated that reducing the transmission rate (assuming a congestion condition) will not improve transmission quality when the actual problem is a fading condition in the wireless component of the communication channel. Despite such limitations, conventional transport layer protocols are often employed to provide the limited (often text-based) Internet content available to wireless devices. As a result, the wireless Internet access user experience is often disappointing.

Another limitation associated with such heterogeneous networks is error control in a scalable wireless communication channel. Conventional wireless error control schemes often employ a simple threshold, wherein if the BER of a packet exceeds the threshold, the packet is dropped. The assumption is that a couple of lost packets during the course of verbal communication will not be too offensive to the users of the system. If the number of dropped packets exceeds a certain threshold, the system may take corrective action, e.g., attempt to transition to a more reliable communication channel. Alternatively, the system will drop the call

1 and the participants can attempt to reestablish the connection. While these often
2 represent minor (although irritating) inconveniences for the user, they are accepted
3 as a fact of life. When these same assumptions are employed in the data-intensive
4 task of streaming multimedia content over a wireless channel, the minor
5 inconvenience degrades into an unacceptable user experience, thereby limiting
6 adoption of wireless multimedia streaming technology.

7 Thus, what is required as an architecture and related methods for streaming
8 media content over heterogeneous networks, unencumbered by the limitations
9 commonly associated with the transport layer protocol of conventional streaming
10 architectures. Just such a solution is presented in the disclosure to follow.

11 12 **SUMMARY**

13 Described herein are one or more implementations of an invention. Such
14 implementations concern an architecture and related methods for content (e.g.,
15 streaming media) over heterogeneous networks. In accordance with a first
16 example embodiment, a system for streaming media content is presented
17 comprising a server, to provide media content on request through a wireline
18 network, coupled to a wireless host through a network gateway. The network
19 gateway establishes a communication channel from the server to the wireless host
20 through both a wireline network and a wireless network, wherein the
21 communication channel includes a transport layer protocol with control parameters
22 for network elements of each of the wireline network and the wireless network. It
23 will be appreciated, given the description to follow, that incorporation of control
24 parameters associated with each of the wireline and wireless networks enables the
25 network gateway to distinguish transmission problems within each of the

individual networks of the heterogeneous communication channel, facilitating accurate resolution of the transmission problem.

This summary itself is not intended to limit the scope of this patent. Moreover, the title of this patent is not intended to limit the scope of this patent. For a better understanding of the present invention, please see the following detailed description and appending claims, taken in conjunction with the accompanying drawings. The scope of the present invention is pointed out in the appending claims.

BRIEF DESCRIPTION OF THE DRAWINGS

The same reference numbers are used throughout the figures to reference like components and features.

Fig. 1 is a block diagram of a data network within which the teachings of the present invention may be practiced;

Fig. 2 is a block diagram of an example content server suitable for use in the data network, according to one aspect of the present invention;

Fig. 3 is a block diagram of an example wireless host suitable for use in the data network, according to one aspect of the present invention;

Fig. 4 is a block diagram of an example network gateway suitable for use in the data network, according to one aspect of the present invention;

Fig. 5 is a flow chart of an example method for reliably delivering content across heterogeneous networks, according to one embodiment of the present invention;

Fig. 6 is a flow chart of an example method for improving transmission quality in a heterogeneous network, according to one aspect of the present invention;

1 **Fig. 7** is a flow chart of an example method of dynamic, channel-adaptive
2 error control in the heterogeneous network, according to one embodiment of the
3 present invention;

4 **Fig. 8** is a communication diagram of an example dynamic, channel-
5 adaptive error control scheme for scalable media over a wireless network,
6 according to one aspect of the present invention;

7 **Fig. 9** is a block diagram of an example computing system suitable for use
8 in implementing one or more of a content source, a wireless host, and/or a network
9 gateway, according to one example embodiment; and

10 **Fig. 10** is a graphical illustration of an example storage medium including
11 instructions which, when executed, implement the teachings of the present
12 invention, according to one embodiment of the present invention.

13 14 **DETAILED DESCRIPTION**

15 This invention concerns an architecture and related methods for streaming
16 content over a heterogeneous network including a wired network component and a
17 wireless network component. In this regard, a transport layer protocol for reliably
18 delivering content over a heterogeneous network is introduced. According to one
19 aspect of the invention, the heterogeneous network transport layer protocol enables
20 network elements (e.g., a network gateway) to distinguish wireless network
21 transmission problems (e.g., fading, multipath, etc.) from wireline transmission
22 problems (e.g., network congestion), to facilitate a focused and timely resolution to
23 such identified problems. In addition, a dynamic, channel-adaptive error control
24 scheme for scalable media is introduced, further enhancing error control in the
25 wireless network component of the heterogeneous network architecture.

As disclosed herein, the invention is described in the broad general context of computing systems of a heterogeneous network executing program modules to perform one or more tasks. Generally, these program modules include routines, programs, objects, components, data structures, etc. that perform particular tasks or implement particular abstract data types. In this case, the program modules may well be included within the operating system or basic input/output system (BIOS) of a computing system to facilitate the streaming of media content through heterogeneous network elements.

As used herein, the working definition of computing system is quite broad, as the teachings of the present invention may well be advantageously applied to a number of electronic appliances including, but not limited to, hand-held devices, communication devices, KIOSKs, personal digital assistants, multiprocessor systems, microprocessor-based or programmable consumer electronics, network PCs, minicomputers, mainframe computers, wired network elements (routers, hubs, switches, etc.), wireless network elements (e.g., basestations, switches, control centers), and the like. It is noted, however, that modification to the architecture and methods described herein may well be made without deviating from spirit and scope of the present invention.

Example Heterogeneous Network

Turning to **Fig. 1**, a block diagram of an example heterogeneous network is presented within which the teachings of the present invention may be practiced, according to one example embodiment. More specifically, Fig. 1 illustrates a block diagram wherein one or more wireline hosts 102, 104 (e.g., content servers) are coupled to provide data to one or more wireless hosts 118, 120 through a heterogeneous network comprised of a wired network component 106 and wireless

1 network component 114 and 116. Unlike conventional heterogeneous data
2 networks, however, Fig. 1 includes an innovative network gateway 110, 112 at the
3 edge between the wired link and the wireless link (note, not necessarily between
4 the wired network component and the wireless network component as the wired
5 network component may well be coupled to control elements of the wireless
6 network component in accordance with conventional wireline protocols). As will
7 be discussed more fully below, network gateway 110, 112 is responsive to an
8 innovative heterogeneous network transport layer protocol to dynamically adjust
9 one or more transmission control parameters of one or more of the wired network
10 component 106 and/or wireless network component 114/ 116 to alleviate
11 transmission problems identified therein. That is, network gateway 110, 112 in
12 conjunction with the innovative heterogeneous network transport layer protocol
13 facilitate an efficient and accurate transmission control across heterogeneous
14 network boundaries, thereby improving transmission quality and performance as
15 compared to conventional solutions utilizing disparate transport layer protocols.

16 Network gateway 110, 112 is an intermediate network node coupling the
17 wireline communication link with a wireless communication link. In this regard,
18 the network gateway may well be a router, a switch, a hub, a wireless basestation
19 controller, and the like. As introduced above, the gateway 110, 112 utilizes an
20 innovative heterogeneous network transport layer protocol to accurately and
21 efficiently distinguish wireline-related transmission problems from wireless-
22 related transmission problems. In this regard, gateway 110, 112 computes and/or
23 receives information in the transport layer regarding one or more transmission-
24 related parameters, and issues control parameters uniquely associated with
25 problems which arise in the individual network components (106, 114/116). For
example, upon receiving an indication in the transport layer of a fading condition

1 in the wireless link, gateway 110, 112 initiates corrective action in the wireless
2 link. Similarly, if congestion is detected in the wireline communication link,
3 gateway 110, 112 initiates corrective action designed to alleviate the problem in
4 wireline link. That is, gateway 110, 112 does not initiate corrective action based
5 merely on assumptions of where a transmission problem may lie, but rather on
6 evidence of such a problem gathered from the heterogeneous network transport
7 layer protocol.

8 As used herein, wireline host 102, 104 are each intended to represent any of
9 a wide variety of computing devices which provide content to requesting users.
10 According to one implementation, one or more of wireline host 102, 104 is a
11 content server, to stream media content to requesting users upon request. In this
12 regard, hosts 102, 104 may well comprise a personal computing system, a server
13 computing system, a media server farm, a KIOSK, thin client hosts, thick client
14 hosts, and the like. According to one implementation, to be described more fully
15 below, wireline host 102, 104 invokes an instance of a content delivery application
16 upon receiving a request for content from a requesting user. The wireline host
17 102, 104 implements congestion control at the transport layer based, at least in part
18 on congestion feedback received from the gateway. Error control is performed at
19 the application layer (e.g., by the content delivery application) and/or at the
20 transport layer based, at least in part, on information received from the mobile host
21 and gateway.

22 Wireless hosts 118, 120 are also intended to represent any of a wide variety
23 of computing devices with wireless communication facilities. In this regard,
24 wireless hosts 118, 120 may well comprise cellular telephones, digital wireless
25 telephones, personal digital assistant (PDA) with wireless communication
facilities, a personal computing system with wireless communication facilities, and

1 the like. As will be developed more fully below, wireless host 118, 120 invokes an
2 instance of an application to request and receive content from a wireline host 102,
3 104. According to one aspect of the invention, wireless host 118, 120 identifies
4 transmission problems (e.g., multipath, fading, high BER problems, etc.) in the
5 wireless communication channel (e.g., 114, 116), and informs the gateway of such
6 wireless transmission problems via the innovative heterogeneous transport layer
7 protocol described herein. As with the wireline host (102, 104), the wireless host
8 118, 120 implements application level and/or transport layer error control based, at
9 least in part, on information received from the gateway.

10 The wired 106 and wireless 112/114 networks are each intended to
11 represent a wide variety of such networks known in the art. In this regard, the
12 wired network 106, for example, may well be comprised of a local area network
13 (LAN), wide-area network (WAN), private network, global public network
14 (Internet), and the like. Similarly, the wireless network 112/114 may well be
15 comprised of a cellular telephony network, a third generation digital
16 communication system network, a personal communication system (PCS) network,
17 a digital cellular telephony network, a two-way paging network, a two-way radio
18 network, a one-way broadcast radio network, a wireless local area network
19 (WLAN) and the like. Similarly, the wireless communication channel between
20 114/116 and 118/120 is intended to represent any of a wide variety of wireless
21 communication links such as, for example, a radio frequency (RF) communication
22 link, an infrared (IR) communication link, and the like commonly associated with
23 any of the wireless communication networks above. Insofar as each of these
24 disparate network architectures were designed to address different problems and
25 operating environments, it is readily understood that the transport layer protocol of
a wireline network often differs from that of a wireless network. This gives rise to

1 the control problems introduced above when one attempts to stream media across
2 such network boundaries. The innovative network gateway 110, 112 overcomes
3 the problems with the prior art in traversing such disparate network boundaries.

4 **Fig. 2** is a block diagram of an example content server 200 suitable for use
5 in the data network as, for example, wireline host 102, 104, according to one
6 example embodiment. In accordance with the illustrated example implementation
7 of Fig. 2, server 200 is generally comprised of control logic 202, a system memory
8 204, one or more applications, and a media component 208. As used herein, server
9 200 is communicatively coupled to a wireline network (e.g., 106) to provide a
10 requesting user with media content. According to one implementation, an
11 application 206 (e.g., streaming application) is selectively invoked to retrieve
12 media content from some source (e.g., file, audio/video device, audio/video tape,
13 etc.) into a local memory 204 for encoding and transmission to a requesting user
14 via media component 208. In this regard, server 200 is intended to represent any
15 of a wide variety of network servers for streaming media content. For purposes of
16 clarification from the onset it should be noted that, as presented in Figs. 2, 3, 4 and
17 9, thin grey lines denotes a control flow of information, while the wider, darker
18 lines denote the flow of data content (e.g., media content).

19 Control logic 202 selectively invokes and controls various functional
20 elements of content server in response to requests for content. According to one
21 embodiment, control logic 202 receives a request for media content from a remote
22 host (e.g., wireless host 118), and selectively invokes an instance of a content
23 delivery application 206 (e.g., a media streaming application) along with the
24 resources of media component 208 to satisfy the request for content. According to
25 one implementation, media component 208 is one of a plurality of applications 206
available on content server 200.

1 As used herein, control logic 202 and system memory 204 are intended to
2 represent any of a wide variety of such devices known in the art. In this regard,
3 control logic 202 may well include one or more of a processor, a microcontroller,
4 an application specific integrated circuit (ASIC), programmable logic device
5 (PLD), programmable array logic (PAL), and/or instructions which, when executed
6 by one of the foregoing devices, implements such control logic. Similarly,
7 memory 204 is intended to represent any of a wide variety of volatile and/or non-
8 volatile memory such as, for example, random access memory, read-only memory,
9 a hard disk, an optical disk, a magnetic tape, and the like.

10 As introduced above, media component 208 is selectively invoked by
11 control logic 202 in response to a request for content from server 200. As shown,
12 media component 208 is comprised of a congestion controller 212, an application
13 layer error control interface 216, a partial checksum generator 214 and a
14 transmission rate controller, each communicatively coupled as depicted.

15 Congestion controller 212 analyzes heterogeneous network transport layer
16 protocol information for an indication from a coupled gateway (110, 112) of
17 congestion in the wireline communication link coupling server 200 to the gateway
18 (110, 112). More particularly, congestion controller 212 monitors the
19 heterogeneous network transport layer protocol for an indication in the congestion
20 control parameter of congestion in the wireline communication link. According to
21 one implementation, the congestion control packet indicates the packet loss caused
22 by congestion on the wireline communication link. Upon receiving such a control
23 parameter, congestion control monitor calculates the available wireline
24 communication channel bandwidth, e.g., based on the packet loss ratio and round-
25 trip transmission time. For additional background information in this area, please
see U.S. Provisional Application S/N 60/208,659 titled "Resource Allocation in

Multi-Stream IP Network for Optimized Quality of Service” and filed on 5/31/2000.

Additionally, congestion monitor 212 may also determine whether a period of retransmission of the dropped packets has elapsed, i.e., based at least in part on the available wireline communication channel bandwidth. Factoring in one or more of such congestion characteristics, congestion controller 212 instructs transmission rate controller 210 to smoothly adjust the transmission rate of the streaming content in an attempt to alleviate the congestion problem in the wireline communication link. Moreover, if the dropped packets are still available in local memory 204, and congestion controller 212 determines that there is available bandwidth and the retransmission time-out has not lapsed, congestion controller 212 may effect re-transmission of one or more of the dropped packets. According to one implementation, congestion controller 212 also takes into account one or more wireless communication link control parameters (e.g., fading, high BER indications, etc.) when making a determination of whether to retransmit dropped packets.

As alluded to above, transmission rate controller 210 controls the rate at which requested content is delivered to the wireline communication link for delivery to the requesting host (e.g., wireless host 118). As depicted, transmission rate controller 210 receives the content from the content delivery application (206) via an error control interface 216 of the application and a partial checksum generator 214. Based, at least in part, on instructions received from the congestion controller 212 and the partial checksum generator 214, transmission rate controller 210 delivers the requested content to the wireline communication link to match the available bandwidth of the wireline communication link.

1 The partial checksum generator 214 calculates a checksum of at least a
2 portion of a subset of packets to be sent to the wireline link for purposes of error
3 control in the transport layer. Conventional transport layer protocols (e.g., TCP,
4 UDP, etc.) will drop any packet identified with errors. This is due, in part, to the
5 relatively low BER experienced in wireline network communications. On the
6 contrary, wireless communication channels suffer from a relatively high BER. As
7 a consequence, wireless applications have been designed with robust error control
8 mechanisms that enable them to be much more error tolerant. Recognizing this
9 error tolerance attribute of wireless applications in general, and multimedia
10 applications in particular, the heterogeneous network transport layer protocol
11 utilizes a partial checksum policy.

12 According to one implementation of the partial checksum policy, the partial
13 checksum generator 214 generates a checksum for only a portion of a given
14 packet. According to one implementation, the portion of the packet for which a
15 checksum is calculated is user-defined. As used herein, the checksum is only
16 applied to the packet header, or a portion thereof. According to the partial
17 checksum policy introduced herein, errors identified in the partial checksum
18 portion of the packet are subject to transport layer error control facilities, to be
19 described more fully below. Errors identified in the non-partial checksum portion
20 of the packet are subject to the error control facilities of the application (i.e.,
21 application layer error control). In this regard, the heterogeneous network
22 transport layer protocol provides for higher throughput taking into account the
23 error control facilities of the wireline and wireless communication links, as well as
24 the application layer.

25 The error control interface 216 monitors the heterogeneous network
transport layer protocol for high bit-error rate indications from the gateway (110,

1 112). According to one implementation, if error control interface 216 receives a
2 control parameter indication of a high BER, server 200 strengthens its error control
3 facilities, and vice versa. Note that the error control is performed at the
4 application level, because different applications will usually have different
5 tolerances for error. According to an alternate implementation, media component
6 208 utilizes error control at the transport layer based, at least in part, on the content
7 delivery application and the degree of BER, as denoted in the control parameter.
8 According to one implementation, the indication of a high bit-error rate (BER) is
9 determined from a negative acknowledgement (NACK) received from the
10 requesting user, i.e., the wireless host (118, 120).

11 **Fig. 3** is a block diagram of an example wireless host 300 suitable for use in
12 the data network, according to one aspect of the present invention. As used herein,
13 wireless host 300 may well be used in data network 100 as wireless host 118, 120.
14 As introduced above, wireless host 300 is intended to represent a wide range of
15 computing systems with wireless communication capability such as, for example,
16 wireless telephony devices, PDA's with wireless communication capability, one-
17 and/or two-way pagers, personal computing systems, and the like.

18 In accordance with the illustrated example embodiment of Fig. 3, wireless
19 host 300 is generally comprised of control logic 302, memory 304, application(s)
20 306 and a media component 308, each coupled as shown. As above, each of
21 control logic 302 and memory 304 are intended to represent such logic and
22 memory as are typically found on such devices, and need not be described further.
23 According to one example implementation, applications 306 include an application
24 for receiving and rendering content from a communicatively coupled server (e.g.,
25 102, 104). According to one example, applications 306 include an Internet

1 browser application that enables a wireless host to access and receive content (e.g.,
2 media content) from an Internet server (e.g., 102, 104).

3 When an application is invoked to access and receive content from a
4 wireline server (102, 104), control logic 302 invokes an instance of media
5 component 308 to provide the requisite interface at the transport layer. In
6 accordance with the illustrated example implementation of Fig. 3, media
7 component 308 includes a fading time-out monitor 310, a header analyzer 312, a
8 bit-error rate controller 314 and an application error correction interface 316, each
9 coupled as shown.

10 In accordance with the partial checksum policy of the heterogeneous
11 network transport layer protocol, header analyzer 312 checks the correctness of the
12 packet header. If a packet header is corrupted, the wireless host sends a negative
13 acknowledgment (NACK) to the gateway (110, 112). BER monitor 314 issues one
14 or more control parameters denoting the BER in the transport layer protocol, in an
15 attempt to invoke error control facilities of either the application layer and/or the
16 transport layer. As introduced above, and will be described more fully below, upon
17 receipt of a NACK, the gateway (110, 112) determines whether to resend the
18 corrupt packet, to issue a command to the sender to strengthen error control (i.e., at
19 one or more of the transport or application layer), etc. Often, this decision is
20 based, in part, on a number of indicators factored together and is not, typically, the
21 result of one single indicator (e.g., the NACK alone).

22 The fading time-out monitor 310 utilizes a time running at the wireless host
23 during the connection, wherein every received packet resets the timer. If the timer
24 runs out, the wireless host determines that a fading condition is occurring. In
25 response, as soon as the wireless host emerges from the fading condition, fading
time-out monitor 310 issues a fading control parameter, via the heterogeneous

1 network transport layer protocol, to the gateway (110, 112) notifying the gateway
2 of the fading condition. According to one implementation, fading time-out
3 monitor 310 includes an indication of the last packet received before fading and
4 the first packet received after fading in the fading control parameter.

5 In addition to the foregoing, media component 308 includes a BER monitor
6 314, which is in charge of reporting a BER to the sender at a regular interval.
7 According to one implementation, BER monitor 314 calculates the number of
8 corrupted packet headers during a period of time to establish a BER control
9 parameter. In accordance with equation (1), below, the packet error probability
10 (Pe) is expressed as a quotient of the number of corrupt packets (i) divided by the
11 number of packets transmitted during a specified period of time.

$$12 \quad Pe = \frac{i}{n}. \quad (1)$$

13
14
15 Taking the error dependency into consideration, the relationship between BER and
16 Pe may well be expressed as:

$$17 \quad Pe = \sum_{i=1}^{hl} b^i (1-b)^{hl-1}, \quad (2)$$

18
19
20 where: b is a single bit error;

21 hl is the length of a packet header; and

22 Pe is the packet error probability.

23 Once computed, BER monitor 314 issues a BER report, via a BER control
24 parameter in the transport layer protocol to at least the gateway and, in certain
25 implementations, to the content server via the gateway.

1 **Fig. 4** is a block diagram of an example network gateway 400 suitable for
2 use in the data network, according to one aspect of the present invention. As used
3 herein, gateway 400 may well be used in data network 100 as gateway 110 and/or
4 gateway 112, to implement the teachings of the present invention. As introduced
5 above, gateway 400 implements an innovative heterogeneous network transport
6 layer protocol to distinguish wireless transmission problems from those
7 transmission problems occurring in the wireline network component of the
8 heterogeneous network. Additionally, as introduced above and described more
9 fully below, the heterogeneous network transport layer protocol includes one or
10 more control parameters which enable the gateway 400 to efficiently and
11 accurately address transmission problems occurring in either network component.

12 As shown, gateway 400 is depicted comprising control logic 402, memory
13 404, one or more applications, and a media component 408, each coupled as
14 shown. As introduced above, gateway 400 is intended to represent a wide range of
15 computing systems. In this regard, control logic 402 and memory 404 are each
16 intended to represent any of a wide variety of control logic and memory which are
17 typically found in such computing systems, and need not be further described
18 herein.

19 In accordance with the illustrated example embodiment of Fig. 4, control
20 logic 402 invokes an instance of the gateway media component 408 to facilitate the
21 heterogeneous network transport layer protocol between network components of a
22 heterogeneous network. As shown, gateway media component 408 includes a
23 congestion monitor 410, data buffers 412, a weighted scheduling function 414, and
24 one or more retransmission modules such as, for example, the NACK
25 retransmission module 416 and the fading retransmission module 418, each
coupled as depicted.

In general, data is received from a content provider (e.g., wireline host 102, 104) through a wireline network component into buffers 412 for subsequent delivery to the requesting wireless host (e.g., 118, 120) via a communication channel in the wireless network component. The buffer caches all data from the content server (200) in the order received. While retransmission is needed, or fading happens, the delivery time and loss ratio can be reduced by just transmitting the packet in the wireless link. That is, by buffering a certain amount of content locally, within the gateway, gateway buffers 412 reduce the burden on the heterogeneous network of having to re-transmit data from the originating content server. According to one implementation, the size of the gateway buffers is designed to be large enough to accommodate the packets sent during a typical fading condition. According to one implementation, the size of the gateway buffer is ten times (10x) the size of the number of packets received during an average fading condition.

Congestion monitor 410 calculates the amount of congestion experienced in the wireline network component of the heterogeneous network, and issues a report on the congestion on a periodic basis (or, upon request). As introduced above, congestion monitor 410 may well utilize any of a number of measures of congestion known in the art such as, for example, the packet loss ratio over a period of time, the round trip transmission time, and the like.

Suppose the whole data can be represented as a binary time series, $\{x_i\}_{i=1}^n$, where x_i takes 1 if the i th packet has arrived successfully and 0 if it is lost. Denoted $\hat{p} = n_{01}/n_0$ and $\hat{q} = n_{10}/n_1$, where n_{01} is the number of times in the observed time series when 1 follows 0 and n_{10} is the number of times when 0 follows 1. n_0 is the number of 0s and n_1 is the number of 1s in the trace. The packet loss ratio is given by

$$P_L = \frac{\hat{q}}{\hat{p} + \hat{q}}$$

According to one implementation, the available bandwidth is estimated in the content server media component (e.g., 208) based on the periodic issuance of congestion control parameters by gateway 400. Considering the delivery requirements of the content to be delivery (e.g., multimedia), the content server media component smoothly adjusts the transmission rate based, at least in part, on the adjustment interval, packet loss ratio, and the bandwidth variations.

NACK retransmission module 416 monitors the heterogeneous network transport layer protocol for NACK (negative acknowledgements) issued by the wireless host to denote that the wireless host has not accurately received a packet. According to one example implementation, the wireless host only issues a NACK if a delay constraint can be satisfied, i.e., if the delay caused by retransmitting the packet does not exceed a threshold. According to one implementation, the sequence number of the corrupted packet is sent with the NACK. On receiving the NACK, the NACK retransmission module 416 controls gateway buffer 412 to retransmit the packet if it can be identified within the buffer.

Similarly, fading retransmission module 418 monitors the transport layer protocol for end of fading notifications from the wireless host. As alluded to above, the gateway is unable to sense the onset of a fading condition when it happens and continues to broadcast data packets to the wireless host as though everything were fine. Once the wireless host emerges from the fading condition, it issues a fading control parameter to the fading retransmission module 418 to inform the gateway of the fading condition. As alluded to above, the fading control parameter includes an indication of the last received packet before fading,

1 from which fading retransmission module 418 determines what packet with which
2 to restart the transmission. In the meanwhile, the fading control parameter includes
3 an indication of the first received packet after fading, from which fading
4 retransmission module 418 determines which packets needed to be retransmitted.
5 According to one implementation, fading retransmission module 418 attempts to
6 restart transmission with the first packet sent during the fading condition, if it is
7 still available in the gateway buffers 412.

8 Weighted scheduling function 414 controls the transmission of content from
9 the gateway buffers 412 to the receiving wireless host (300) based on a number of
10 factors, e.g., content type, state of the wireless host, volatility of the content to be
11 delivered, time sensitivity of the content to be delivered, etc. Considering the
12 content priorities, the weighted scheduling function can be adapted to control the
13 buffer so that optimal end-to-end quality can be achieved. Representative
14 examples of weighted schedulers are Weighted Fair Queuing (WFQ) and Virtual
15 Clock (VC). If the weights (the rates) in WFQ (in VC) corresponding to the
16 individual queues are equal, the algorithms divide the capacity of the output link
17 emulating a time-division multiplexer (TDM), i.e. they send the packets out almost
18 in the same order that TDM would. If the weights are not equal, the queues share
19 the capacity according to their weights, for example 1:1:3:4. If any of the queues
20 has not enough packets to send out, the other queues share its portion according to
21 their weights. WFQ and VC are implemented using virtual time and timestamps
22 which determine the order of transmission.

23 **Example Transport Layer Protocol Elements and Data Structures**

24 As introduced above, the innovative heterogeneous network transport layer
25 protocol is designed to enable network elements (e.g., the gateway) to efficiently

1 and accurately distinguish transmission problems occurring in any of a plurality of
2 network components, e.g., a wireless network component, a wireline network
3 component, and the like. In this regard, as introduced above, the heterogeneous
4 network transport layer protocol includes a plurality of control parameters which,
5 when enabled, provide an indication to one or more network elements of the
6 existence of an error condition in one or more of the network components. In
7 addition, one or more of the network elements (e.g., the sender, receiver, gateway,
8 etc.) are responsive to receipt of such control parameters to attempt to resolve
9 problems denoted by the received control parameters.

10 In accordance with the illustrated example implementation introduced
11 above, four (4) example heterogeneous network transport layer protocol control
12 parameters are introduced, which enable a gateway to accurately identify the
13 source of transmission problems, and to effect a network component-specific
14 resolution to the identified problem. In this regard, as introduced above, the four
15 heterogeneous network transport layer protocol control parameters include:

16 ***Congestion control parameter:*** issued by the gateway to one or more
17 wireline component network elements to throttle the delivery of content in
18 accordance with current congestion statistics;

19 ***BER control parameter:*** issued by the wireless host to periodically inform
20 the gateway and/or one or more wireline component network elements of the
21 current BER at the wireless host;

22 ***Fading control parameter:*** issued by the wireless host to inform the
23 gateway that the wireless host has just emerged from a fading condition; and

24 ***Negative Acknowledge (NACK) control parameter:*** issued by the wireless
25 host to the gateway to inform the gateway that the last packet sent was dropped
due to high bit errors.

1 In accordance with the illustrated example introduced above, each of these
2 control parameters include information which enable the receiving network
3 element to act on the information to correct the problem identified. It is to be
4 appreciated, given the foregoing discussion, that the innovative heterogeneous
5 network transport layer protocol could well utilize more or less control parameters
6 to facilitate reliable content delivery through disparate network components of a
7 heterogeneous network without deviating from the spirit and/or scope of the
8 present invention. Indeed, an expanded set of control parameters which address
9 additional transmission issues is anticipated within the scope and spirit of the
10 present invention.

11 12 **Example Operation and Implementation**

13 Having introduced the functional and architectural elements of the
14 heterogeneous data network 100, an example operation and implementation is
15 developed with reference to Figs. 5 through 9. For ease of illustration, and not
16 limitation, the teachings of the present invention will be illustrated within the
17 exemplary implementation of a media streaming application introduced above. In
18 accordance with this example implementation, a content source (wireline server)
19 responds to a content request from a wireless host by streaming content through
20 the heterogeneous network (wireline/wireless components) to the host (e.g.,
21 cellphone, PDA, etc.) via the gateway and a wireless communication channel.
22 Unlike conventional heterogeneous content delivery solutions, however, the
23 implementation described herein utilizes an innovative heterogeneous network
24 transport layer protocol which enables one or more network elements (e.g., the
25 gateway) to distinguish wireline transport problems from wireless transport

1 problems. Accordingly, the discussion to follow will continue to reference
2 elements of Figs. 1-4.

3 **Fig. 5** is a flow chart of an example method for delivering content across
4 heterogeneous networks, according to one embodiment of the present invention.
5 As shown, the method of Fig. 5 begins with block 602, wherein wireline server
6 102 receives a request for content from a wireless host 118. More particularly,
7 wireless host 118 issues a request for content via wireless communication channel
8 114, gateway 110, wireless network 108, wireline network 106 to host 102. As
9 introduced above, upon receiving such a request, control logic 202 of host 102
10 invokes an instance of media component 208 to facilitate content delivery.

11 In block 604, host 102 identifies one or more characteristics of the network
12 through which the request was received. According to one implementation, host
13 102 identifies whether the requesting user is coupled to the content server 102
14 through a homogeneous network or a heterogeneous network, block 606. If the
15 request was received through a homogeneous network (e.g., just a wireline
16 network), content delivery server 102 fulfills the request utilizing a standard
17 network transport protocol (e.g., TCP, UDP, etc.) and error protection facilities.

18 If, however, the request was received over a heterogeneous network
19 consisting of, for example, a wireline network component and a wireless network
20 component, the process continues utilizing the resources of the heterogeneous
21 network transport layer protocol to satisfy the request, block 610. In this regard, as
22 described above, each of the elements participating in the delivery of content to the
23 requesting user (e.g., server, gateway, wireless host) issues and responds to one or
24 more control parameters enabling the gateway 110 to monitor and uniquely control
25 transmission characteristics in one or more of the network components comprising
the heterogeneous network until the transmission is complete, block 612.

1 If, during the course of transmission, one of the elements identifies an
2 unacceptable transmission condition 614, it is reported to the gateway 110. In
3 response, gateway 110 distinguishes transmission problems arising in a wireline
4 network component of the heterogeneous network from those occurring in the
5 wireless network component based on the control parameter received, and
6 selectively invokes a network component-specific remedy to alleviate the
7 transmission problem, block 616. For example, receipt of a fading control
8 parameter from the wireless host is uniquely identified by gateway 400, which then
9 efficiently addresses the fading problem itself and/or in combination with other
10 problems denoted by other network-centric control parameters.

11 In block 618, gateway 400 determines whether the network-centric remedy
12 initiated in block 616 has worked to resolve the transmission problem. More
13 particularly, in accordance with one embodiment, gateway 400 waits a period of
14 time as it continues to monitor the heterogeneous network transport layer protocol
15 of the communication channel for indications from one or more of the control
16 parameters that the identified condition is subsiding. If so, the process continues
17 with block 612 wherein transmission continues until the desired content has been
18 delivered.

19 If, in block 618 it is determined that the implemented solution to the
20 problem identified in block 614 is not resolving the problem, gateway 400
21 selectively invokes another network-specific remedy to improve transmission, as
22 the process returns to block 616. That is, although the innovative heterogeneous
23 network transport layer protocol enables network elements to efficiently identify
24 transmission problems and implement network-centric solutions, often one or more
25 failure modes may well have multiple causes. For example, a high BER may be
realized as a mobile is entering a fading and/or multipath condition, causing one or

more of the gateway 400, the source 200, etc. to initiate corrective action. Later, when the BER doesn't improve (and, perhaps, upon receipt of additional control parameters providing additional information into the root cause of the problem), the elements may well attempt additional corrective action measures to improve the transmission quality.

Fig. 6 is a flow chart of an example method for improving transmission quality in a heterogeneous network, according to one aspect of the present invention. In accordance with the illustrated example of Fig. 6, a more detailed description of identifying the source of transmission problems (block 616 of Fig. 5) is presented. As shown, the method begins with block 712 wherein a determination is made as to which network component of the heterogeneous network the source of the transmission problem lies. As introduced above, the heterogeneous network transport layer protocol includes a plurality of control parameters, issued by one or more network elements, to identify network-centric transmission problems. Thus, a preliminary determination is made in block 704 of whether the transmission problem lies in the wireless or wireline link. As introduced above, this determination is made based, at least in part, on the type of control parameter received.

If, in block 704, the gateway 400 identifies a congestion condition exists in the wireline network component (104) of the heterogeneous network 100, gateway 400 issues a congestion control parameter to upstream network elements supporting the wireline communication channel for corrective action. At the source, the server invokes congestion control measures in an attempt to remove the identified congestion from the wireline link, block 706.

If, in block 704, gateway 400 receives a control parameter from the wireless host, the type of control parameter received will determine the corrective action

1 invoked by the gateway and, perhaps, other network elements supporting the
2 heterogeneous communication channel. Accordingly, if in block 708 gateway 400
3 receives a negative acknowledgment (NACK) from the wireless host (300),
4 gateway (400) determines whether the delay constraint can be satisfied and, if the
5 packet remains in the buffer, retransmits the packet to the wireless host (300).

6 If, in block 710, gateway 400 receives an end-of-fading (EoF) control
7 parameter from the wireless host (300), gateway 400 retransmits the packets in the
8 gateway buffer (412), starting with the packet sequence number identified in the
9 EoF notification.

10 If, in block 704, gateway 400 receives a high BER control parameter from
11 the wireless host (300), gateway 400 forwards the received control parameter to
12 the content server (200) which may well implement (or strengthen) error control
13 mechanisms at the application layer. The source (e.g., server (200)) may invoke an
14 innovative channel-adaptive error control scheme for scalable media over a
15 wireless network component, block 712. As will be described more fully below,
16 the channel-adaptive error control mechanism reduces high BER conditions by (1)
17 accurate channel estimation that predicts or estimates in near-real time the bit-error
18 rate or frame erasure rate of the communication channel; (2) given the estimated
19 BER and/or FER, the gateway invokes a channel-adaptive hybrid delay constrained
20 Automatic Repeat reQuest (ARQ) and Forward Error Correction (FEC) error
21 control scheme; and (3) modifies bit allocation between a base content layer and an
22 enhanced content layer at the source to limit the increase in wireless channel power
23 required to improve the received BER at the wireless host.

24 Those skilled in the art will appreciate that gateway 400 in association with
25 the heterogeneous network transport layer protocol unite transmission control of
once disparate networks to facilitate reliable content delivery over such a

1 heterogeneous network. Without such a heterogeneous network transport layer
2 protocol the delivery of, for example, media content would be nearly impossible as
3 the occurrence of transmission problems in either network component would result
4 in both network components implementing corrective action, the result of which is
5 typically the rapid degradation of content throughput, without a commensurate
6 improvement in goodput – i.e., an unacceptable user experience.

7 Turning next to **Fig. 7**, a flow chart of an example method of dynamic,
8 channel-adaptive error control in the heterogeneous network is presented,
9 according to one embodiment of the present invention. As introduced above, many
10 media streaming applications benefit from a sophisticated error control scheme
11 which makes them relatively error resilient. Such error control schemes do not,
12 however, anticipate transmission over a heterogeneous network much less a
13 network consisting of a wireless network component. Those skilled in the art will
14 appreciate that often, error correction in a wireless communication channel
15 requires an increase in transmission power (effectively raising the noise floor in
16 the wireless network). Moreover, not all applications benefit from such an
17 integrated error control system. Thus, despite the fact that a certain element of
18 error control identified above is relegated to the application layer, it remains
19 important to implement some level of error control in the transport layer.

20 In accordance with but an example illustrated embodiment, the method of
21 Fig. 7 begins with block 802 wherein the gateway 400 generates an estimate of
22 communication channel status based, at least in part, on one or more channel
23 quality metrics and statistical modeling techniques, block 802. According to one
24 implementation, channel modeling is performed by error control modules 413/216.

Channel Modeling

To accurately estimate channel status and have it mean anything in an error control scheme, one or more of the gateway 400 and or the source (e.g., server 200) monitors several channel-related characteristics on a near real-time basis. More particularly, each of gateway 400 and source 200 are depicted comprising error control modules 413 and 216, respectively. It is to be appreciated that such error control modules need not be identical but may, in certain implementations, include cross-over functionality. Thus, in describing the channel modeling feature, one or more of the error control modules 413/216 may perform this function.

According to one implementation, a channel bit error rate (BER) and fading depth characteristics are fed to the error control module 413/216 from the wireless host (300) via the physical layer, and applied to a dynamically evolving channel status model generated by error control module 413/216. According to one implementation, aiming at simulating wireless channels characterized by a slow-, highly correlated fading characteristics, a moderately slow-motion model of wireless host (300) is used by error control module 413/216 in its channel simulation. It is to be appreciated that the teachings of the error control module 413/216 may readily be applied to utilize other wireless host fading models, or that one of a plurality of models be dynamically chosen based, at least in part, on the monitored channel characteristics identified above.

According to one example implementation, the channel model employed by error control module 413/216 multiplies signal transmitted from the encoder by a Rayleigh-distributed amplitude factor taking into account the effect of fading, to which is added a Gaussian white noise signal. Error control module 413/216 generates the sequence of Rayleigh amplitude values by summing two squared

Gaussian random variables, e.g., $(x(i), y(i))$ and taking the square root of the result. Channel correlation is taken into account by applying a low pass filter to the sequence of Gaussian values before squaring and summing them. The low pass filter used to account for channel correlation is built assuming the speed of the wireless host (300). Given such a speed, error control module 413/216 calculates the Doppler frequency, f_{dopp} , of the channel in accordance with equation (3) below:

$$f_{dopp} = \frac{v_{mob}}{c_0} f_0 \quad (3)$$

where: v_{mob} is the speed of the mobile host;

c_0 is the speed of light; and

f_0 is the carrier frequency of the wireless channel.

It can be shown that for a broad range of parameters, the sequence of data-block success and failure could itself be approximated by means of a two-state Markov chain, whose transition probability matrix is given by:

$$Mc = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix} \quad (4)$$

where: p and $1-q$ are the probabilities that the j^{th} block transmission is successful, given that the $(j-1)^{\text{th}}$ block transmission was successful or a failure, respectively.

Using this model, the steady-state transport block error rate, P_{BL} , is given by

$$P_{BL} = \frac{1-p}{2-p-q} \quad (5)$$

For a Rayleigh fading channel with fading margin F , the average transport block error rate (P_{BL}) and the Markov parameter (q) can be expressed as

$$P_{BL} = 1 - e^{-1/F}, \quad (6)$$

$$\text{and } q = 1 - \frac{Q(\theta, \rho\theta) - Q(\rho\theta, \theta)}{e^{1/F} - 1}, \quad (7)$$

$$\text{where } \theta = \sqrt{\frac{2/F}{1 - \rho^2}}. \quad (8)$$

In the above, $\rho = J_0(2\pi f_{dopp}T)$ is the correlation coefficient of two successive samples (spaced by T seconds, which equals to 10 ms, 20 ms, 40 ms, or 80 ms in 3G network) of the complex Gaussian fading channel with Doppler frequency, f_{dopp} , which calculates in eqn. (3), $J_0(\cdot)$ is the Bessel function of the first kind and zero order, and $Q(\cdot, \cdot)$ is the Marcum-Q function given by

$$Q(x, y) = \int_y^\infty e^{-\frac{(x^2+w^2)}{2}} I_0(xw) w dw. \quad (9)$$

Continuing with the process of Fig. 7, having generated a statistical model of the communication channel, error control module 413/216 employs one or more error control schemes to reduce the distortion experienced in the receiver, block 804. More specifically, according to one example embodiment, error control module 413/216 selectively invokes and dynamically modifies a hybrid ARQ and FEC error control scheme, while being sensitive to delay constraints based, at least in part, on the modeled channel status.

Channel-Adaptive, Delay Constrained Error Control

Those skilled in the art will appreciate that, considering the time varying channel conditions, the distortion to be experienced in the receiver consists primarily of two parts: (1) the source distortion caused by media encoding, and (2) channel distortion caused by transmission errors. According to one implementation, error control module 413/216 derives a rate-distortion (R-D) of the scalable media, then formulates the distortion associated with the forward error

control (FEC), the automatic repeat request (ARQ), and a hybrid combination of the two in order to reduce the distortion perceived by the receiver.

Rate-Distortion of the Scalable Media

As introduced above, the distortion in the receiver consists of source distortion and channel distortion. Regardless of transmission errors, source distortion is determined in large part by the operational R-D function $SourceD(R)$ of the media encoder, where R is the source rate. On the other hand, bit error rate leads to channel distortion, $Distortion(n)$, which depends primarily on the error resilience capability of the media codec. If variable length coding (VLC) is used in the media stream, a single lost codeword may well cause severe channel distortion. For video media, another important issue is interframe error propagation because errors may be visibly perceptible to the end-user over many consecutive frames. The distortion at the receiver can thus be modeled as:

$$D(R_s) = SourceD(R_s) + \sum_{n=1}^{R_s} P(n) Distortion(n) \quad (10)$$

where: $P(n)$ is the probability that the n^{th} bit is corrupted while the related bits are error free.

Rate-Distortion of the Forward Error Control (FEC)

Those skilled in the art will appreciate that forward error control (FEC) is well suited for error control in communication applications. However, the varying channel conditions commonly associated with wireless communication channels limits its effective application to wireless communication applications, since the worst-case design scenario often requires a prohibitively large amount of overhead to effectively use FEC. Leveraging the channel modeling above, however, error

control module 413/216 employs adaptive FEC to facilitate delivery of media content over a wireless communication channel.

According to one implementation, error control module 413/216 applies different error correction rates based, at least in part, on the priority class of the content. According to one embodiment, the higher the priority, the more error correction rate is applied. On the other hand, each layer error control rate is also determined by the wireless network conditions. If the network conditions are good, error control module 413/216 reduces the rate of FEC. While the network conditions are bad, error control module 413/216 increases the FEC rate. Accordingly, the protection of the media is adaptive to media data priority as well as the current network conditions. In this regard, error control module may well employ FEC to the media stream while reducing the total amount of overhead typically associated with FEC in wireless communication channel applications. According to one implementation, the rate-distortion model introduced above is modified with the FEC as shown in equation (11), below:

$$D(R_S) = SourceD(R_S) + \min_{(N_m, K_m)_{m=1, M}} \left\{ \sum_{m=1}^M (w_m \times \sum_{l=1}^{L_m} (D_{error}(l, m) \times P(l, m))) \right\} \quad (11)$$

where: w_m is the distortion weight for the m^{th} layer;

D_{error} represents the channel distortion that l^{th} symbol in the m^{th} layer is corrupted;

M is the total number of layers; and

L_m is the total number of symbols in the m^{th} layer.

To avoid overlapped channel distortion, $P(l,m)$ is the probability that the l^{th} symbol of the m^{th} layer is corrupted, while the corresponding symbols in the previous layers are correct. N_m, K_m represent block length and information length, respectively, of layer m , respectively, which are related to the probability of $P(l,m)$.

Rate-Distortion of Automatic Repeat reQuest (ARQ)

Closed loop error control techniques such as ARQ have been shown to be more effective in certain situations and implementations than FEC. However, retransmission of corrupted data frames introduces additional delay into the communication channel, delay which may prove prohibitively expensive in near-real time content delivery applications. Accordingly, the content delivery system 100 implements a hybrid ARQ architecture, wherein the sender monitors certain transmission channel characteristics, along with a protection degree. According to one example, content source 200 monitors a delay constraint, $D_{\text{constrained}}$, which is limited by media frame rate, current roundtrip transmission time (RTT) and the estimated time consumed by processing procedure, $D_{\text{processing}}$. In addition, according to one implementation, the sender also designates the protection degree for the content to be delivered to the requestor. Accordingly, the maximum transmission times of current packet N_{max} is calculated as:

$$N_{\text{max}} = \frac{D_{\text{constrained}} - D_{\text{processing}}}{RTT} \quad (12)$$

The receiver (e.g., wireless host 300), when any error is detected after channel decoding, decides whether to issue a retransmission request based on the delay bound of the packet. If the delay bound has not expired, the request will be sent. In response, the content source (200) will only transmit corresponding information of the error packet. It can be observed, given the foregoing, that the

receiver has more redundant information using ARQ. If the redundant information is organized elaborately, the probability of error can be reduced considerably as N_{\max} increases. Thus, the rate-distortion model of ARQ is formulated as:

$$R = R_S + R_1 + \sum_{i=2}^{N_k} P(i) * R_i \quad (13)$$

$$D(R) = \text{Source}D(R_S) + P(N_{\max}) * D(N_{\max}) \quad (14)$$

where: $P(i)$ is the probability of the i^{th} transmission of the error packet;

$D(N_k)$ is the distortion of the N_k^{th} transmission of packet k ;

R_i is the bit rate for the i^{th} transmission.

Rate-Distortion of a Hybrid FEC and ARQ scheme

As introduced above, FEC and ARQ are both useful means of error control. FEC is suited for real time communications under high BER and short fading wireless channel, while ARQ does well under low BER and long fading wireless channel conditions. As introduced above, a hybrid ARQ scheme may achieve both delay bound and rate effectiveness. The procedure of the hybrid FEC and ARQ is similar to the ARQ scheme described above, with the difference being that in each retransmission stronger protection is added at the content server (200), wherein the higher protection is achieved and the overhead maintained at a relatively small size. Thus, in accordance with one example implementation, the rate-distortion model of equation (11) above is modified as:

$$D(R_S) = \text{Source}D(R_S) + \min_{(N_m, K_m) m=1, M} \left\{ \sum_{m=1}^M (w_m \times \sum_{l=1}^{L_m} (D_{\text{error}}(l, m) \times P(l, m, N_{\max}^m))) \right\} \quad (15)$$

That is, equation (11) is modified with the inclusion of $P(l, m, N_{\max}^m)$, which is the probability that the l^{th} symbol of the m^{th} layer at N_{\max}^m retransmission

is corrupted while the corresponding symbols in the previous layers are correct.

Continuing with block 806 of Fig. 7, once the appropriate ARQ and/or FEC error correction scheme is selectively invoked (block 804), the sender (200) monitors bit allocation among and between data layers to reduce end-to-end distortion, while maintaining a minimal amount of overhead (i.e., directly related to power consumption). One of the challenging tasks for streaming multimedia content, for example, over a wireless channel is how to allocate the bit budget to the source and the channel. Because the media content serviced is scalable, it can be divided into two parts, e.g., important media data (IMD) and less important media data (LMD). According to one implementation, the error processing scheme of processing system 100 applies different error control schemes to IMD and LMD according to their respective priorities, and characteristics of the media (e.g., audio vs. video vs. text, etc.).

Bit Allocation Scheme

Let $R(t)$ represent the channel rate available for transmission at time t . Let $R_s(t)$, $R_{Hybrid}(t)$ and $R_{FEC}(t)$ denote the source rate, hybrid error control rate for IMD and FEX rate for LMD at time t , respectively. Then the problem becomes to allocate the available bit rate at time t such that the optimal $R_s(t)$, $R_{Hybrid}(t)$ and $R_{FEC}(t)$ are achieved by minimizing the end-to-end distortion under the constraint:

$$R_s(t) + R_{Hybrid}(t) + R_{FEC}(t) \leq R(t) \quad (16)$$

Using the rate-distortion models developed above, equation (16) can be re-written as:

$$(R_s, R_{Hybrid}, R_{FEC})_{opt} = \arg \min_{R_s, R_{Hybrid}, R_{FEC}} \{SourceD(R_s) + P(N_{max}) * D_{IMD}(N_{max}) + \min_{(N_m, K_m)_{m=1, M}} \sum_m (w_m \times \sum_{l=1}^{L_m} D_{error}(l, m) \times P(l, m))\} \quad (17)$$

subject to $R_s + R_{Hybrid} + R_{FEC} \leq R$.

1
2 The second term on the right side of equation (17) represents the channel distortion
3 resulting in IMD with error correction rate R_{hybrid} , while the third term denotes the
4 channel distortion resulting in LMD with error correction rate R_{FEC} . In this regard,
5 the bit allocation scheme presented takes into account the relative priority of
6 content within each layer and the channel conditions to reduce distortion and
7 improve transmission quality, minimizing any increase in transmitted overhead
8 and, thus, power consumption.

9 In block 808, a determination of whether transmission has completed is
10 made. If so, the process ends, otherwise it continues with block 802, as described
11 above.

12 To further illustrate the process introduced in Fig. 8, above, attention is now
13 drawn to **Fig. 8** wherein a graphical illustration of an example dynamic, channel-
14 adaptive error control scheme for scalable media over a wireless network is
15 presented, according to one aspect of the present invention. As above, thick black
16 lines in the figure denote data content, while narrower grey lines denote control
17 content.

18 In accordance with the illustrated example implementation of Fig. 8, the
19 communication model 900 is presented comprising a scalable content encoder
20 (e.g., media encoder) 904, channel-adaptive, unequal error control module 906, the
21 wireless communication channel 910 (e.g., 114), a channel decoder 912, and a
22 scalable source content decoder 918. In accordance with the error control scheme
23 introduced above, the scalable source content encoder 904, which may reside in
24 the source (200) and/or gateway (400), receives content 902 and encodes it in
25 accordance with the compression mechanism associated with the wireless
communication channel 910. Once encoded, the encoded content is passed

1 through the channel-adaptive, priority sensitive error control module 906 (again, at
2 source (200) and/or gateway (400)) before communication to a requesting wireless
3 host (300) via the wireless communication channel 910.

4 Once received at the wireless host (300), the content is passed through a
5 channel decoder 912 before passing through a scalable source content decoder 916
6 to retrieve a representation of the content sent from the source (200). In
7 accordance with the teachings of the present invention, the wireless host (300)
8 determines whether the base layer (e.g., comprising the IMD) error threshold has
9 been met, and whether the delay constraint has been met, block 914. If not, such
10 an error condition is issued (e.g., in a high BER control parameter) to the gateway
11 (400) and/or the source (200). In addition, the channel decoder 912 provides
12 feedback information (e.g., high BER, NACK, fading condition control
13 parameters) associated with the quality of the transmission over the wireless
14 communication channel to the source (200) and/or gateway (300).

15 As introduced above, the channel estimation module 928 in the gateway
16 receives such information and dynamically generates a model of the wireless
17 communication channel. As introduced above, the error control module 906 may
18 well reside in the source (200) (i.e., as application error control interface). In
19 response to such information, module 930 controls a bit allocation among the
20 source content and channel protection to reduce the distortion of the received
21 content, while limiting the amount of overhead added to the communication
22 channel, thereby limiting any additional power consumption incurred as a result of
23 the error control mechanism.

Example Computing Environment

Fig. 9 illustrates an example of a suitable computing environment 1000 on which one or more elements of the innovative heterogeneous network for streaming media content may be practiced. It should be appreciated that computing environment 1000 is only one example of a suitable computing environment and is not intended to suggest any limitation as to the scope of use or functionality of the streaming architecture. Neither should the computing environment 1000 be interpreted as having any dependency or requirement relating to any one or combination of components illustrated in the exemplary computing environment 1000.

The example computing system 1000 is operational with numerous other general purpose or special purpose computing system environments or configurations. Examples of well known computing systems, environments, and/or configurations that may well benefit from the heterogeneous network transport layer protocol and dynamic, channel-adaptive error control schemes described herein include, but are not limited to, personal computers, server computers, thin clients, thick clients, hand-held or laptop devices, multiprocessor systems, microprocessor-based systems, set top boxes, programmable consumer electronics, wireless communication devices, wireline communication devices, network PCs, minicomputers, mainframe computers, distributed computing environments that include any of the above systems or devices, and the like.

Certain features supporting the streaming architecture, e.g., the heterogeneous network transport layer protocol, the dynamic, channel-adaptive error control scheme, etc. may well be described in the general context of computer-executable instructions, such as program modules, being executed by a computer. Generally, program modules include routines, programs, objects,

1 components, data structures, etc. that perform particular tasks or implement
2 particular abstract data types.

3 As shown in Fig. 9, the computing environment 1000 includes a general-
4 purpose computing device in the form of a computer 1002. The components of
5 computer 1002 may include, but are not limited to, one or more processors or
6 execution units 1004, a system memory 1006, and a bus 1008 that couples various
7 system components including the system memory 1006 to the processor 1004.

8 As shown, system memory 1006 includes computer readable media in the
9 form of volatile memory 1010, such as random access memory (RAM), and/or
10 non-volatile memory 1012, such as read only memory (ROM).

11 While not depicted, computer 1002 may further include other
12 removable/non-removable, volatile/non-volatile computer storage media. By way
13 of example only, a hard disk drive, a magnetic disk drive (e.g., a "floppy disk"),
14 and/or an optical disk drive may also be implemented on computing system 1002
15 without deviating from the scope of the invention. Moreover, it should be
16 appreciated by those skilled in the art that other types of computer readable media
17 which can store data that is accessible by a computer, such as magnetic cassettes,
18 flash memory cards, digital video disks, random access memories (RAMs), read
19 only memories (ROM), and the like, may also be used in the exemplary operating
20 environment.

21 Bus 1008 is intended to represent one or more of any of several types of bus
22 structures, including a memory bus or memory controller, a peripheral bus, an
23 accelerated graphics port, and a processor or local bus using any of a variety of bus
24 architectures. By way of example, and not limitation, such architectures include
25 Industry Standard Architecture (ISA) bus, Micro Channel Architecture (MCA)
bus, Enhanced ISA (EISA) bus, Video Electronics Standards Association (VESA)

1 local bus, and Peripheral Component Interconnects (PCI) bus also known as
2 Mezzanine bus.

3 A user may enter commands and information into computer 1002 through
4 input devices such as keyboard 1034 and/or a pointing device (such as a “mouse”)
5 1036 via an input/output interface(s) 1040. Other input devices 1038 may include
6 a microphone, joystick, game pad, satellite dish, serial port, scanner, or the like,
7 coupled to bus 1008 via input/output (I/O) interface(s) 1040.

8 Display device 1042 is intended to represent any of a number of display
9 devices known in the art. A monitor or other type of display device 1042 is
10 typically connected to bus 1008 via an interface, such as a video adapter 1044. In
11 addition to the monitor, certain computer systems may well include other
12 peripheral output devices such as speakers (not shown) and printers 1046, which
13 may be connected through output peripheral interface(s) 1040.

14 Computer 1002 may operate in a networked environment using logical
15 connections to one or more remote computers via one or more I/O interface(s)
16 1040 and/or network interface(s) 1054.

18 **Alternate Embodiments - Computer Readable Media**

19 Turning to **Fig. 10**, an implementation of one or more elements of the
20 architecture and related methods for streaming content across heterogeneous
21 network elements may be stored on, or transmitted across, some form of computer
22 readable media in the form of computer executable instructions. According to one
23 implementation, for example, instructions 1102 which when executed implement
24 one or more elements of the heterogeneous network transport layer protocol, the
25 media component of the content source 200, wireless host 300, and/or the network
gateway 400, and/or instructions 1104 to implement the dynamic, channel adaptive

error control scheme of the content delivery system 100 may well be embodied in computer-executable instructions. As used herein, computer readable media can be any available media that can be accessed by a computer. By way of example, and not limitation, computer readable media may comprise “computer storage media” and “communications media.”

As used herein, “computer storage media” include volatile and non-volatile, removable and non-removable media implemented in any method or technology for storage of information such as computer readable instructions, data structures, program modules, or other data. Computer storage media includes, but is not limited to, RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to store the desired information and which can be accessed by a computer.

“Communication media” typically embodies computer readable instructions, data structures, program modules, or other data in a modulated data signal, such as carrier wave or other transport mechanism. Communication media also includes any information delivery media.

The term “modulated data signal” means a signal that has one or more of its characteristics set or changed in such a manner as to encode information in the signal. By way of example, and not limitation, communication media includes wired media such as a wired network or direct-wired connection, and wireless media such as acoustic, RF, infrared, and other wireless media. Combinations of any of the above are also included within the scope of computer readable media.

Fig. 10 is a block diagram of a storage medium 1100 having stored thereon a plurality of instructions including instructions to implement a heterogeneous

1 network transport layer protocol 1102 and/or a dynamic, channel-adaptive error
2 control scheme for scalable media according to yet another embodiment of the
3 present invention.

4 As used herein, storage medium 1100 is intended to represent any of a
5 number of storage devices and/or storage media known to those skilled in the art
6 such as, for example, volatile memory devices, non-volatile memory devices,
7 magnetic storage media, optical storage media, and the like. Similarly, the
8 executable instructions are intended to reflect any of a number of software
9 languages known in the art such as, for example, C++, Visual Basic, Hypertext
10 Markup Language (HTML), Java, eXtensible Markup Language (XML), and the
11 like. Accordingly, the software implementation of Fig. 10 is to be regarded as
12 illustrative, as alternate storage media and software embodiments are anticipated
13 within the spirit and scope of the present invention.

14 Although the invention has been described in language specific to structural
15 features and/or methodological steps, it is to be understood that the invention
16 defined in the appended claims is not necessarily limited to the specific features or
17 steps described. It will be appreciated, given the foregoing, that the teachings of
18 the present invention extend beyond the illustrative example implementation
19 presented above. For example, the teachings of the present invention are equally
20 applicable to an implementation wherein the content server is the wireless host and
21 the receiving host is an element of a wireline network. Similarly, the
22 heterogeneous network transport layer protocol and/or the dynamic, channel-
23 adaptive error control scheme may also be beneficially applied to streaming
24 content solely through a wireless network. Thus, the specific features and steps are
25 disclosed as merely example forms of implementing the claimed invention.